Original Article: Empirical

Research Ethics
2019, Vol. 15(2) 1–21

DOI: 10.1177/1747016118820497
journals.sagepub.com/home/rea
SAGE

Research Ethics in Correspondence Testing: An Update

Eva Zschirnt
European University Institute, Italy

Abstract

Correspondence testing to research discrimination in the marketplace has become common and the use of internet applications has allowed researchers to send greater numbers of applications. While questions of research ethics always arise when planning a correspondence test, the issue receives relatively little attention in published correspondence tests. This paper addresses the question of ethics in correspondence testing in the age of ready internet access. It focusses on the ethical issues that arise in correspondence testing, looking at potential problems (regarding voluntary participation, informed consent, deception, entrapment of employers, employers' rights) and possible solutions, and technical challenges. European country examples show that the ethical questions raised in correspondence testing have to be renegotiated depending on the national context. The paper argues that correspondence testing, if planned carefully and executed responsibly, can meet most of the ethical requirements of Social Science ethics guidelines.

Keywords

research ethics, correspondence testing, discrimination, field experiments

Open discrimination has decreased with the adoption of anti-discrimination legislation. Yet, discrimination continues to occur in more subtle and hidden ways. Field experiments in the marketplace, such as audit and correspondence studies,

Corresponding author:
Eva Zschirnt, European University Institute, Max Weber Programme, Via dei Roccettini 9, 50014, San Domenico di Fiesole, Italy.
Email: eva.zschirnt@eui.eu

provide important information on the extent of systematic differential treatment in the labour market and are currently seen as the best way to measure discrimination. The methodology of correspondence testing has existed since the late 1960s (Jowell and Prescott-Clarke, 1970) and correspondence tests relying on fictitious candidates have been carried out most prominently in the labour and housing market on various grounds of discrimination, such as that of ethnicity/race, gender, disability, or sexual orientation.[1] Questions of research ethics arise in the planning stages of such experiments, because correspondence testing relies on covert research where participants are not aware that they are part of an experiment. While this violates core research principles such as informed consent and voluntary participation that have been enshrined in ethical guidelines across disciplines, sociological research ethics guides argue that covert research can be justified under certain strict provisions.

Despite the growing number of correspondence tests on various grounds of discrimination, ethical questions are rarely thoroughly addressed in published correspondence tests on ethnic discrimination in hiring decisions. Most authors only refer to Riach and Rich (2004a), who discuss the deceptive nature of field experiments in detail, and there are rarely any references to Banton (1997), who focusses on the rejection of Swedish research proposals for a correspondence test on ethical grounds. Similarly, contributions on research ethics in the US context by Edley Jr. (1993), Fix et al. (1993), or Pager (2007) are also hardly mentioned. The teaching case presented by Connor (2000) that describes ethical challenges and arguments against a proposed testing that were voiced by an Internal Review Board (IRB) is also largely ignored. Even in more recent publications on the methodology, such as the book edited by Gaddis (2018a), questions of research ethics are not discussed: Lahey and Beasley (2018) only very briefly mention ethical issues related to the amount of correspondence that researchers send to employers, while Crabtree (2018) explicitly states that getting IRB approval is not discussed in his chapter on the steps included in conducting audit studies.

Does this mean that a consensus over the legitimacy of using correspondence testing has emerged or, is it rather the case that most researchers focus more thoroughly on the questions of ethics in the preparation of their experiments, but do not include further information in the publication of their results? Almost 15 years have passed since the publication of Riach and Rich's article and since then correspondence testing has become more widespread and diverse.[2] Furthermore, researchers increasingly use the internet to find vacancies and send applications. This greater number of CVs being sent out at relatively low costs is a great opportunity for researchers, because it allows them to test numerous new variables or combinations of variables that require larger samples and to obtain more nuanced results. However, it also means that more employers are affected by a correspondence test. It is therefore worthwhile to revisit the question of ethics, to account for

recent technological developments, and to address challenges that the computerisation of field experiments has created.

This paper investigates ethical concerns involved in correspondence testing and argues that the thorough preparatory work required in the planning stages of a correspondence test can mitigate most ethical concerns frequently voiced by opponents of the technique. Looking at (mostly) sociological research ethics guidelines, experiments on discrimination have been acknowledged as instances where conducting covert research can be justified under strict circumstances. The fact that correspondence tests have been approved by ethical commissions[3] in numerous countries shows that this methodology has been recognised as, so far, the best way to measure discrimination in hiring decisions and that ethical concerns can be minimised.

Correspondence testing: An introduction to the technique

Field experiments, of which correspondence tests are only one method, allow researchers to observe behaviour in real-life situations. One of the earliest field experiments was conducted by LaPiere (1934), who travelled with two Chinese friends through the USA and documented if they were accepted as guests in hotels or restaurants. After half a year had passed, he contacted these establishments again and asked if they would accept Chinese customers. Although most establishments had previously welcomed them, almost all expressed their refusal to do so in a written questionnaire. This shows that certain behaviour cannot be detected by simply asking people 'How would you react?', but that it can be observed in real-life situations and that considerable differences between the outcomes can occur.

Since the 1960s, field experiments have been used to study the phenomenon of discrimination in the labour market, making use of in-person audits and written correspondence testing (Cherry and Bendick, 2018; Gaddis, 2018b; Zschirnt, 2016). In recent years, and in European countries in particular, correspondence testing has been deemed to be one of the most suitable methods to identify and measure discrimination in the labour market and in hiring decisions in particular (Schneider et al., 2014: 14). In a correspondence test, researchers apply in writing to real-life vacancies and present potential employers with (at least) two substantially equal and thus interchangeable candidates, who differ only in the characteristic to be studied, such as ethnicity. The companies included in the experiment are not named and the exclusive focus on anonymised aggregated data guarantees the protection of participants' privacy.

Correspondence tests are carefully planned experiments and most published experiments contain a detailed research design section. These careful considerations before the actual testing takes place show that correspondence tests require a

lot of time and detailed preparation before they can be conducted (for a detailed discussion of the methodology, see Gaddis, 2018a). Once this labour-intensive preparatory work has been done, Bendick and Nunes point out that correspondence testing is an 'innovative research technique […] that offers laboratory-like controlled conditions in quasi-experiments in real-world hiring situations' (2012: 238). The preparatory work should, however, not be limited to the research design, but also include the inevitably arising discussion of ethical issues. Correspondence tests constitute an ethical challenge, especially concerning the responsibilities of the researcher towards the research participants. It is therefore not surprising that correspondence tests are often met with scepticism regarding their compliance with research ethics standards, and in particular the criteria of informed and voluntary consent and the use of deception.

Ethical issues in correspondence testing

Researching hiring discrimination using correspondence tests lies at the intersection of sociology and economic research, and developments in the fields of research ethics in the Social Sciences also influence researchers planning correspondence tests. The most frequently voiced ethical objections are now addressed in detail.

Objection: Correspondence testing infringes the principles of voluntary participation and informed consent

Two of the fundamental ethical principles underlying research across disciplines are that 'potential research subjects should be given the opportunity to refuse participating in research' (Dench et al., 2004: 56) and that they make this decision based on comprehensive and accurate information (63). The insistence on voluntary participation and informed consent can be traced back to medical experiments conducted by the Nazi regime and the subsequent development of the Nuremberg Code of 1947, which above all emphasises the importance of voluntary and informed consent of research participants (Israel, 2015: 27).[4] This emphasis on voluntary and informed consent has been extended far beyond medical experiments and is also applied in Social Science research.

Research ethics guidelines both on the supranational level, such as the non-binding EU Code of Ethics (Dench et al., 2004), and on the national level, emphasise the importance of ensuring voluntary participation and informed consent of research subjects since 'the consent requirement is intended to prevent invasions of personal integrity' (National Committees for Research Ethics in Norway (NESH), 2006: 13). However, sociological research ethics guidelines recognise that there are exceptions, where research would not be possible if voluntary and informed consent had to be obtained first.

> In certain cases, participant's freedom and self-determination can be respected even though consent has not been obtained beforehand. […], exceptions […] can be made in certain cases in situations in which the research does not imply physical contact with the research subjects, where the data being processed is not particular sensitive, and *where the utility value of the research clearly exceeds any disadvantages that might be inflicted on the subjects*. (NESH, 2006: 14, emphasis added)[5]

It lies in the nature of correspondence testing that employers cannot make a voluntary choice to participate in a scientific experiment and it is not possible to inform them of their participation and the goal of the study. Starting with Bovenkerk in the 1990s, researchers have recognised that not informing the research subjects conflicted with their right to provide or refuse their consent (1992: 33). Yet, according to Bovenkerk, three reasons justify breaking the principle of informed consent: first, hiring decisions are not a private matter, and hiring discrimination is unlawful; second, if field experiments are carefully prepared and carried out, there is almost no detrimental effect on the employers tested; and third, it is normal hiring decisions that are observed, and researchers 'do not lure employers into a situation in which they are enticed to deviate from their normal course of action' (Bovenkerk, 1992: 33–34).

Breaking the principle of informed consent has also been addressed by other researchers who argued that breaking it 'is a crucial feature of this type of research, as informing participants would invalidate the experiment' (Blommaert et al., 2014: 964). This has also been recognised in national laws. In the case of Sweden, Bursell (2007) refers to Swedish law, which states that 'research without the participant's informed consent can still be carried out' if the research meets certain conditions, such as being 'of high societal importance' (Bursell, 2007: 9). Looking at the USA, Pager (2007) also refers to legal provisions on conducting research without obtaining informed consent:

> …a human subjects institutional review board…'may … waive … informed consent provided (1) the research involves no more than minimal risk to human subjects; (2) the waiver or alteration will not adversely affect the rights and welfare of the subjects; (3) the research could not practicably be carried out without the waiver or alteration; and (4) whenever appropriate, the subjects will be provided with additional information after participation'. Each of these conditions can arguably be satisfied in the context of audit studies of discrimination. (Pager, 2007: 126)

While Pager explicitly refers to audit studies, these conditions are also met in correspondence tests of discrimination. The methodology of correspondence testing is usually judged admissible under certain (strict) provisions due to the higher societal interest to measure discrimination, even if it breaks the principles of voluntary participation and informed consent.

In order to mitigate the potential negative effects of breaking the principle of informed consent, the British Sociological Association proposes that this consent

could be obtained post-hoc (2002: 4). While Midtbøen (2014) decided to contact only some of his unknowing research participants to recruit participants for follow-up interviews, I am only aware of one study in which all participants were systematically informed post-hoc (Liebkind et al., 2016). However, Pager (2007) argued that 'for human resource personnel or managers who are thought to be discriminating, the consequences may be more serious than if no attention were brought to the audit whatsoever' (Pager, 2007: 127). The research team of the Expert Council of German Foundations on Integration and Migration decided against informing their participants. They reasoned that informing employers post-hoc would not improve the chances of minority applicants in the future and that it might pose a problem for further research by making the technique too well-known. It should still be possible for other researchers to conduct correspondence tests without too large an awareness of the methodology among employers. Following Pager's argument, the Expert Council of German Foundations on Integration and Migration's research team claimed that obtaining post-hoc consent could potentially prove problematic for the employees responsible for the hiring decisions. Thus, in order to limit these potential damages incurred by individuals, it can be argued against seeking post-hoc consent. Furthermore, informing unknowing participants post-hoc will also take more time away from the human resources personnel. Finally, informing participants post-hoc could give them the possibility to try to sue researchers or the ethical approval bodies and researchers should also avoid causing harm to themselves or their colleagues.

Another option would be informing employers via the media that correspondence testing will take or has taken place. Yet, Swedish research showed that employers did not change their hiring behaviour after extensive media coverage of a correspondence testing had occurred (Carlsson and Rooth, 2012).

Objection: Researchers are deceiving their research participants

The above-mentioned principle of informed consent includes that participants make their choice based on accurate information. But are there situations in which it is acceptable to deceive participants? Dench et al. argue in the EU Code of Ethics that 'there are varying debates about whether deception is ever acceptable' and that 'the conclusions vary depending on the methodological, philosophical and moral stance' (2004: 68) of the researchers.

Correspondence testing relies on the deception of research subjects, because employers are presented with fictitious applicants who pretend to be real candidates. This deceptive nature of field experiments has been the focus of Riach and Rich's (2004a) seminal article, where it featured prominently in the title 'Deceptive field experiment – Are they ethical?'. They start their argument saying that field experiments 'constitute an unequivocal procedure for charting, over the time, the

effectiveness, or otherwise, of equal opportunity legislation' (2004a: 458). They then turn towards the context in which the deception of employers occurs. Using Bovenkerk's argument that the action performed by the researcher is 'a non-genuine transaction performed in a manner which is not infrequent in the labour market' (Bovenkerk, 1992: 34), Riach and Rich elaborate on the notion that testing 'takes place in an arena where deception is a regular and acknowledged activity' (2004a: 461). They justify the deception of employers because

> …a lack of veracity is endemic in these markets; […] great harm is done to the social fabric by discriminatory practices in such markets; […] minimal inconvenience is imposed on the entrepreneurs in the experiment, and […] the technique provides evidence with a degree of accuracy and transparency which is not available from any other procedure (Riach and Rich, 2004a: 463).

Similarly, Edley argued that the use of testing was justified, because 'the moral costs of deception are outweighed by the great benefit of developing a clearer understanding of the social disease' (Edley, 1993: 378). The deceptive nature of field experiments is thus seen as necessary to obtain information about the socially harmful practice of discrimination.

One of the strongest arguments of why the use of deception in research may be justified is its 'resemblance to an accepted method for gathering evidence for the enforcement of anti-discrimination law' (Banton, 1997: 416), which many courts, including the US Supreme Court, have endorsed (Banton, 1997; Pager, 2007). While some methodological differences exist between testing for research where employers are only sampled once and testing for law enforcement purposes, where one employer is tested multiple times (Pager, 2007), the methodology was initially developed in the 1960s/1970s in response to the adoption of anti-discrimination laws, to monitor their effectiveness (Pager and Western, 2012). Cherry and Bendick (2018) as well as Boggs (1998) provide great overviews of the development of scholarship and activism in the fight against (mostly) housing discrimination in the USA. In 1982, the US Supreme Court strengthened the position of these scholars and activists in Havens Realty Corp vs. Coleman,[6] when it gave its approval for this methodology. Over the years, US courts have confirmed the legal standing of testers, and 'broaden[ed] their endorsement of this methodology' (Pager, 2007: 127). Even though these cases concerned testing for legal reasons and not research, Pager argued that 'implicit in these holdings […] is the belief that the misrepresentation involved in testing is worth the unique benefit this practice can provide in uncovering discrimination and enforcing civil rights laws' (Pager, 2007: 127). Similarly, the endorsement of the methodology shows that deception is seen as regrettable but unavoidable:

> […] we have long recognised that this requirement of deception was a relatively small price to pay to defeat racial discrimination. The evidence provided by testers […] is a major resource in

> society's continuing struggle to eliminate the subtle but deadly poison of racial discrimination. (Boggs et al., 1993: 367)

The legal situations, for example if results obtained by testing are accepted in courts, vary depending on the national context. The case of the USA is the best known, yet evidence obtained through testing is also recognised in discrimination cases in several European countries. While Rorive (2009) provides a good overview of testing in 11 European countries (Belgium, Czech Republic, Denmark, Finland, France, Hungary, Latvia, The Netherlands, Slovakia, the United Kingdom, and Sweden), van der Plancke (2007) and Calvès (2007) focus on the situation in Belgian and French courts, respectively. Researchers often provide the reasoning of the courts next to guidelines by professional scientific organisations to justify the use of audit or correspondence tests and the element of deception included in this methodology when applying for IRB approval.

Objection: Correspondence testing can have negative consequences for employers who unwillingly participated in the experiment

Another principle that researchers should adhere to is the 'principle of no harm'. Research should not harm research participants, researchers themselves, or future researchers. Opponents of correspondence tests have objected that researchers try to trap employers and catch them in unlawful behaviour, that employers suffer a loss of time by being included in an experiment, that employers' privacy is being breached and that an employer's reputation might suffer from the unwilling participation in a correspondence test.

The first argument brought forward is that researchers try to trap employers if the experiment encourages research subjects to behave illegally. This argument and the fear that researchers might be held liable for such an entrapment has already been addressed in the 1990s by Bovenkerk (1992) and Edley Jr. (1993). According to Bovenkerk, this 'concern is ill-conceived as discriminating employers break the legal rules probably more than the researcher does' (1992: 34). Most importantly, researchers only observe normal hiring practices; they do not lure employers into a trap of acting in a way that they would not have under different circumstances.

Second, opponents of correspondence testing argue that employers suffer from a loss of time by assessing fictitious applicants. This argument has, for example, been brought forward by the former US house-speaker Newt Gingrich, who argued against funding for the Equal Employment Opportunity Commission, because 'the use of testers […] causes innocent businesses to waste resources' (Gingrich, 1998). Researchers acknowledge that assessing additional fictitious applications may pose a burden on the employers' time (e.g. Pager, 2007; Pager and Western, 2012) and most correspondence tests limit this burden by considering an employer only

once, even if more matching vacancies are published. Finally, invitations to interviews are quickly and politely declined in order to keep the application process as normal as possible for genuine applicants. The loss of time should thus be considered minimal (e.g. Wood et al., 2009). Furthermore, it is assumed that employers do not spent much time on the initial screening of applications.

A third argument used against correspondence testing is the breach of employers' privacy. However, Bovenkerk claims that 'there is no question of breaking legitimate expectations of privacy. Hiring is not an entirely private matter' (1992: 33). He further argues that providing equal opportunities in the hiring process is in the public interest and that discrimination in these public fields has been declared unlawful. This argument is supported by Fix, Galster and Struyk, who point out that the 'behaviours that have been monitored […] involve public, commercial, or professional acts. In most instances there has been a special invitation issued to the public – via a published ad for a job, apartment, or loan' (Fix et al., 1993: 16). Furthermore, researchers only gather data on employers that are publicly available.

Finally, concerns are voiced regarding the reputation of enterprises and possible negative effects of being part of a correspondence test. Pager emphasises that 'efforts must be taken to protect employer identities so that even associations with a study on discrimination cannot be made' (2007: 127). Most studies point out that data is anonymised and only accessible to the core research team. Furthermore, the fact that data is only analysed aggregately also helps to avoid inference about individual employers. Finally, correspondence testing for research is not interested in accusing individual employers of discriminatory behaviour, but in reporting trends in discrimination patterns in a society.

Correspondence testing as an example of covert research

By the nature of the research design, in which employers are not aware that an experiment is being conducted, correspondence tests are an example of covert research. This becomes apparent in the definition provided by the non-binding *EU Code of Ethics for Socio-Economic Research*:

> By definition, covert research means that participation is not voluntary and participants are not able to give informed consent. To some researcher(s,) this is unacceptable. Others argue that, in some circumstances, covert research is the only way in which the necessary information can be collected or difficult situations researched. (Dench et al., 2004: 12)

Numerous professional associations have acknowledged the need for deception when it is absolutely necessary. Information on the conditions under which the use of deception is justifiable are included in the ethical guidelines of, for example, the American Sociological Association (2018), the British Sociological Association (2002), the American Psychological Association (2017), the NESH (2006), the

German Sociological Association (Deutsche Gesellschaft für Soziologie, 2017), and the EU's Respect Project on professional and ethical codes for socio-economic research in the information society (Dench et al., 2004). The British Sociological Association, for example, argues that while 'there are serious ethical and legal issues […,] the use of covert methods may be justified in certain circumstances' (2002: 4). It points out that covert research violates the principle of informed consent and may violate the privacy of research subjects, making it a method that should only be used as a last resort if it is impossible to obtain information using other methods. The *EU Code of Ethics for Socio-Economic Research* also addresses the question of deception in covert research:

> If it is only possible to obtain information through covert research (for example, studies of violent, criminal or subversive groups, or of fraudulent or *discriminatory practices*)[,] how can the researcher balance the need for deception against the value to society of conducting the research? (Dench et al., 2004: 64, emphasis added)

They clearly identify studies on discrimination as one area in which covert research is often the only way to avoid the bias of socially desirable behaviour. Dench et al. even refer to field experiments in the labour market explicitly saying that

> …if a study exploring *discrimination in the recruitment process* involved researchers posing as applicants, informing the recruiters in advance may lead to their acting differently than normal. (Dench et al., 2004: 62, emphasis added)

Using covert research methods is a delicate matter, but, as seen above, it can be justified in situations in which information of a similar quality and richness cannot be obtained using other methodologies.

Matched pair testing versus non-matched pair testing

As was briefly mentioned previously, the number of applications submitted for one vacancy can also have ethical implications that can become important in IRB submissions and discussions of the ethical questions of the research, the most obvious being the time an employer spends on assessing applications (Gaddis, 2018b). However, as Lahey and Beasley (2018) point out, the number of resumes could also affect hiring practices and recruiters' decisions, such as if more very well-qualified applications are received in an on-going hiring process. They emphasise that 'unmatched sets send a less focussed signal and may be less likely to harm a participant's overall view of the market' (Lahey and Beasley, 2018: 91).

While most field experiments have traditionally been designed as matched-pair experiments, some more-recent studies have deviated from the design of matching candidates and only sent out single applications (e.g. Koopmans et al., 2018; Weichselbaumer, 2015, 2016).[7] Researchers using a single application design

often state that they want to minimise the risk of detection of the experiment. However, Riach and Rich (2004b) argue that studies using only a single application per vacancy are tests of 'preferential treatment' in the broader labour market or of a propensity among employers to discriminate rather than actual discrimination, since 'employment discrimination can only occur when an individual employer is confronted with a need to choose' (Riach and Rich, 2005b: 471). Because it is not possible to attribute discriminatory treatment to specific employers, Cherry and Bendick (2018) describe the findings of unpaired audits as 'villainy without villains' (Cherry and Bendick, 2018: 55). So far, the most thorough discussion of matched versus unmatched designs has been provided by Vuolo et al. (2016, 2018), who mainly focus on the statistical implications of this research design choice. The ethical implications of the number of applications sent, is however, not discussed in the literature.

Technical and legal challenges in correspondence testing

Although the ethical issues discussed above are important to obtain ethical approval to conduct a correspondence test, there are also other stages in the planning phase of the experiment that can have ethical or legal implications, such as having to set up contact details using email addresses and phone numbers for the fictitious applicants. To my knowledge, these issues have so far not been addressed in articles on the methodology. Finally, correspondence testing may also require the preparation of photographs, diplomas or work certificates, depending on the national context. This is usually required in German-speaking countries, and will therefore be addressed in the country examples below.

Creating email addresses

Each fictitious applicant requires an email addresses to send applications and receive replies. The email addresses most frequently used in the correspondence tests included in the meta-analysis by Zschirnt and Ruedin (2016) were Gmail.com, Hotmail.com, and Yahoo.com. These providers differ considerably in their terms of services. Yahoo, for example, clearly spells out in Section 3 of its Terms and Services:

> In consideration of your use of the Yahoo Services, you represent that you are of legal age to form a binding contract […]. You also agree to (a) provide true, accurate, current and complete information about yourself […]. (Yahoo, 2012)

Similarly, Microsoft stipulates in its terms of services that 'You agree not to use any false, inaccurate or misleading information' (Microsoft, 2015: Section 4.a.i.).

Furthermore, in its Code of Conduct it emphasises that the account is not to be used for anything illegal and that the account holder is not to 'engage in activity that is false or misleading (e.g., […] impersonating someone else […]' (Section 3.a.i.)). These terms can be potentially problematic for researchers conducting a correspondence test, since it is impossible to 'provide true, accurate, current and complete information' for fictitious applicants.

Google's Terms of Service only state 'Don't misuse our services' (Google, 2014) and do not specifically define who is allowed to open a Google account and which conditions have to be fulfilled. Yet at the very end of the Terms and Services, it is stated that 'The laws of California, U.S.A., […] will apply to any disputes arising out of or relating to these terms or the Services' (Google, 2014). In 2010, the State of California adopted its first online impersonation law – the Senate Bill SB 1411 – which regulates that

> (a) Notwithstanding any other provisions of law, any person who knowingly and without consent credibly impersonates another actual person through or on an Internet Web site or by other electronic means for purposes of harming, intimidating, threatening, or defrauding another person is guilty of a public offense punishable pursuant to subdivision (d). (Simitian, 2010: Section 1)

It is therefore necessary to examine in how far correspondence testing might be considered an impersonation of another *actual* person. Since the fictitious applicants in correspondence tests do not exist in real life, it can be argued that this is not an impersonation of another *actual* person. Furthermore, it should be obvious that correspondence testing is not done for 'the purpose of harming, intimidating, threating or defrauding another person'.

So far, all published correspondence tests I am aware of, with the exception of Neumark et al. (2017) who created their own email provider, have used free and frequently used email providers such as Gmail, Yahoo or Hotmail. To my knowledge, there have never been legal objections to their use.

Next to these legal issues, researchers can also encounter problems with the security settings of free email providers that might limit the possibility to send high numbers of emails using programming scripts. Readers interested in the technical aspects of setting up correspondence tests via email should refer to Crabtree (2018), who devotes an entire book chapter to the issue.

Generating street addresses

Street addresses are another elemental part of the contact details that might become problematic. Eid (2012) used addresses of his research team and colleagues for his Canadian experiment. Wood et al. (2009) decided against such an approach in their UK study out of ethical considerations. They argue that UK employers sometimes carry out background checks, including credit checks. Thus, like most other studies,

they constructed credible fictitious addresses, by using real street names, but non-existing house numbers. The residential areas chosen for the study were based on the ethnic diversity shown in census data (Wood et al., 2009: 23). Another approach used by Bursell (2007) was to use real addresses in residential blocks, but making sure that nobody with a similar name lived there. Similarly, researchers could use addresses of real apartment buildings but provide fake apartment numbers. Although posted responses can be lost, Eid (2012) reported that the grand majority of employers contacted potential candidates by phone, and letters were hardly ever used.

Providing phone numbers

The final element of the contact details is a phone number. Here almost all researchers use the same approach: Phone numbers connected to a voicemail box were set up using mobile phones or online generated phone numbers. The number of phone numbers used varied, however. While Eid (2012) used only two numbers, one for the majority and one for the minority applicant, Wood et al. (2009) had 12 phone numbers, depending on the gender and ethnicity of the fictitious applicant. In all studies the voicemail messages were either standard voicemail messages by the phone provider or recorded without any discernible accent. One of the challenges of using voicemails is matching the response received with the vacancy it was connected to. Furthermore, local legal regulations need to be taken into consideration when it comes to setting up mobile phone accounts (e.g. if a proof of ID is required to open an account).

Managing ethical issues in correspondence testing – European examples

Since this paper has so far predominantly focussed on the theoretical discussion of research ethics in correspondence tests as well as arguments from an ethical perspective pro and contra correspondence testing, the last part of the paper discusses ways in which ethical commissions have dealt with correspondence tests. While most publications of correspondence tests acknowledge the question of research ethics without going into further details, some researchers specifically refer to the ethics bodies and procedures in their countries. Unfortunately, such in-depth information on the ethical approval process is only publicly available for four countries – Sweden, Norway, Finland and Germany.

Sweden, Norway and Finland

To my knowledge, the first country where a correspondence test was stopped by a research ethics commission was Sweden. Swedish researchers wanted to participate in the large International Labour Organisation Project on labour market discrimination in the 1990s; they thus submitted two research proposals using the

correspondence test design outlined by Bovenkerk (1992), but their proposals were not approved by the Swedish Ethics Board. The Board claimed that 'invit[ing] an innocent employer to act in a manner likely to have been made punishable by the time any such research started' (Banton, 1997: 415) posed too big a risk. While a first assessor had not expressed any doubts, a second reviewer concluded that while the research might be in the public interest, the potential consequences for people found guilty of discriminatory behaviour were too big: 'The employer runs both a risk of injury to reputation and a financial risk. It is these risks of injury which so clearly make the proposed experiment ethically unacceptable' (in Banton, 1997: 415). Furthermore, potential issues of liability for the researcher or funding organisations were addressed, and it was recommended that neither researcher should be funded.

As Carlsson and Rooth (2012) point out, the Swedish authorities subsequently reconsidered their position on correspondence testing: 'An important event for this turnaround occurred in 2005 when law students initiated lawsuits against restaurants and night clubs based on situation tests of ethnic discrimination' (Carlsson and Rooth, 2012, 99). Following this change of position, three Swedish correspondence tests were published in 2007 (Attström, 2007; Bursell, 2007; Carlsson and Rooth, 2007). Except for a brief section on ethics in Bursell (2007), the issue of research ethics was not addressed.

The rejection of the Swedish contribution to the ILO Project also affected Norwegian researchers, as Midtbøen (2013) points out:

> Because the method was rejected by the Swedish Council for Social Research, it was assumed that the Research Council of Norway would reach the same conclusion. This is a main reason why Norwegian researchers during the 1990s never even applied for funding of experimental studies of discrimination (Midtbøen, 2014: 52).

Once the Swedish research ethics boards approved the above-mentioned research projects using correspondence testing, researchers in Norway proposed to conduct a field experiment on the Norwegian labour market. Midtbøen (2013) reports that the NESH approved the research design under three strict conditions. First, testing should be conducted in the early phase of the hiring process. Second, the privacy of the individuals in the hiring procedure was to be protected. Third, regarding the recruitment of participants for follow-up interviews, it was emphasised that this should respect the principles of voluntary participation and informed consent.

In 2011, ethical approval was also given for the first Finnish correspondence test, where the guidelines of the Finnish National Advisory Board on Research Ethics also 'lists field experiments in studies about discrimination as an example of a research design where deviating from the principle of informed consent and misleading research subjects is acceptable' (Larja et al., 2012: 142).

Germany

In Germany, so far four correspondence tests on ethnic discrimination in hiring decisions have been conducted (Goldberg et al., 1995; Kaas and Manger, 2012; Schneider et al., 2014; Weichselbaumer, 2016). The Expert Council of German Foundations on Integration and Migration was the only one who extensively addressed the ethical questions. Its report includes a short section on the ethical and legal challenges in correspondence testing (Schneider et al., 2014: 16). It emphasises that the research design was approved by the Ethical Committees of the German Sociological Society and the German Association of Sociologists, as it was judged unproblematic both from a data protection and an ethical point of view. It was argued that the aggregated analysis of the data would not allow inferences about individual employers, and that the use of fictitious applications did not infringe any personal rights (Schneider et al., 2014: 16). Yet, the research team went even further than obtaining ethical approval, and also addressed potential legal problems. While two legal expertise studies by Klose and Kühn (2009, 2010) on the use of correspondence testing had previously been commissioned by the Federal Anti-Discrimination Authority, the Expert Council of German Foundations on Integration and Migration hired these lawyers again to specifically analyse their proposed research design for a correspondence test on the German apprenticeship market (Kühn et al., 2013). Since these expert opinions look at numerous legal concerns raised in regard to correspondence testing, they warrant a more detailed look.

In total, there are now three legal expertise studies on testing available for Germany: the first two expertise studies by Klose and Kühn (2009, 2010) focus on very specific legal questions regarding testing and racial or ethnic discrimination in the area of *'Gewerberecht'* (trade law) (2009), and the use of testing as an instrument in trials regarding the burden of proof in discrimination cases (2010). The third expertise study by Kühn et al. (2013) addresses the legal questions concerning testing as a Social Science research method and focussed explicitly on the Expert Council of German Foundations on Integration and Migration's research design. It is therefore the most relevant publication to be considered here. Regarding criminal law, they focus on the use of certificates or copies thereof, concluding that the testing methodology is protected under the scientific freedom guaranteed by the German Basic Law, and that testing does not fulfil the crime of forgery of documents (e.g. school or university certificates) that are required to submit a complete German application. Furthermore, they claim that researchers do not have to fear being punished for fraud, since testing studies are not intended for unlawful gains of the researchers. Looking at civil law, Kühn et al. argue that claims for liability of the researcher due to the time employers invested in examining a fraudulent application are not likely, since the loss of time is not considered a replaceable damage. Employers are also unlikely to succeed in suing for damages by arguing that the fictitious

applications caused a delay or necessitated a repeated application procedure. Furthermore, Kühn et al. closely look at the German data-protection laws in relation to correspondence testing. According to them, data-protection laws do not apply if the data was anonymised and analysed quantitatively and if no inferences about individuals can be made. The use of publicly available data, such as addresses, is also permitted. This legal expertise thus enabled the researchers of the Expert Council of German Foundations on Integration and Migration to conduct their correspondence test on labour market discrimination.

These individual country examples show that the theoretical concerns regarding the ethical questions in correspondence testing discussed in previous sections of the paper are valid, but can be addressed in well-prepared research designs. In the case of Germany, many of the aforementioned reservations, such as the possibility of committing fraud, of forging documents, of potential damages to employers, or the liability of researchers, have been addressed by legal experts and found not to be an obstacle to conducting a correspondence test. While a similar wealth of information on ethical and legal preparation work was not publicly available for other countries, the examples of Sweden, Norway and Finland show that ethical commissions were quite thorough in their evaluation of the research projects, and eventually decided that a good research design could meet their concerns and that the societal interest to study discrimination was held above the inconveniences that could potentially be caused to an individual employer.

Conclusion

As shown in this paper, breaking core research ethics principles, particularly those of informed consent and voluntary partition, can be justified in the case of correspondence studies on discrimination in the labour market. Using examples from different countries and different ethical committees, it can be seen that researchers were able to obtain ethical approval to conduct correspondence tests if certain strict criteria were met, the most important being to keep the inconvenience to employers at a minimum, to guarantee the confidentiality and privacy of the research subjects, to analyse data in an aggregated form to avoid inferences being made about individual research subjects, and to adhere again to the principles of research ethics in any follow-up research.

Given the rapid growth in the numbers of correspondence tests conducted in recent years, it could be questioned whether more studies measuring discrimination are actually needed. A recent meta-analysis has shown that in the case of ethnic discrimination in hiring, minority candidates have to write on average 50% more applications than equally qualified majority candidates (Zschirnt and Ruedin, 2016). It could therefore also be argued that not studying discrimination would be unethical, as the data provided can help in ‘society’s continuing struggle to

eliminate the subtle but deadly poison of racial discrimination' (Boggs et al., 1993: 367). There is still a need to provide data on the extent of discrimination in hiring decisions to make employers aware of these issues, which may also occur unintentionally, and to lobby for policy changes and stronger and more effective anti-discrimination laws. Considering the strong power imbalance in the hiring process, it is necessary to strengthen the position of applicants to balance the scales. Providing better data on the extent of discrimination of minorities is just a first step in this direction. Using data that was obtained through correspondence testing on the broader labour market to strengthen legal cases against discriminatory employers, as it is already possible in some countries, could be a next step.

Acknowledgements

I thank Arnfinn Midtbøen, Didier Ruedin, Rosita Fibbi and Carolin Fischer, as well as the anonymous reviewer for detailed comments.

Funding

This research was supported by the *nccr - on the move* funded by the Swiss National Science Foundation. All articles in Research Ethics are published as open access. There are no submission charges and no Article Processing Charges as these are fully funded by institutions through Knowledge Unlatched, resulting in no direct charge to authors. For more information about Knowledge Unlatched please see here: http://www.knowledgeunlatched.org.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. An overview of the historical development of field experiments with a focus on ethnic or racial discrimination in the labour market can be found in e.g. Cherry and Bendick (2018), Gaddis (2018b), or Zschirnt (2016).
2. For recent reviews and meta-analyses of correspondence tests, see e.g. Baert (2018), Bertrand and Duflo (2017), Gaddis (2018b), Neumark (2016), Quillian et al. (2017), Rich (2014), or Zschirnt and Ruedin (2016).
3. Unfortunately, information about the composition of ethical commissions or IRBs is often not provided. A rare example is Connor (2000), who gives detailed information about the members of the IRB that rejected the research project and points out the problems with having an all-white review panel deciding on research proposals that deal with racial discrimination.
4. For a historical discussion on the development of research ethics, see: Dingwall (2012), Dench et al. (2004), Hunter (2010), Israel (2015), Nakray (2016), or Wassenaar (2006).
5. The case of Norway is discussed so prominently because it is one of the rare studies where detailed information about the ethical approval process is available.
6. Havens Realty Corp. v. Coleman, 455 U.S. 363 (1982).
7. For an overview on matched and unmatched audit studies, see Table 6.1 in Vuolo et al. (2018).

ORCID iD

Eva Zschirnt https://orcid.org/0000-0002-8918-4146

References

American Psychological Association (2017) Ethical Principles of Psychologists and Code of Conduct. Available at: http://www.apa.org/ethics/code/ethics-code-2017.pdf (accessed 18 August 2018).

American Sociological Association (2018) Code of Ethics. Available at: http://www.asanet.org/sites/default/files/asa_code_of_ethics-june2018.pdf (accessed 18 August 2018).

Attström K (2007) *Discrimination against Native Swedes of Immigrant Origin in Access to Employment*. Geneva: International Labour Organisation.

Baert S (2018) Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. In: Gaddis M (ed) *Audit Studies: Behind the Scenes with Theory, Method and Nuance*. Cham: Springer, pp.63–77.

Banton M (1997) The ethics of practice-testing. *New Community* 23(3): 413–420.

Bendick M and Nunes AP (2012) Developing the research basis for controlling bias in hiring. *Journal of Social Issues* 68(2): 238–262.

Bertrand M and Duflo E (2017) Field experiments on discrimination. In: *Handbook of Economic Field Experiments*. Dordrecht: Elsevier, pp.309–393.

Blommaert L, Coenders M and van Tubergen F (2014) Discrimination of Arabic-named applicants in the Netherlands: An internet-based field experiment examining different phases in online recruitment procedures. *Social Forces* 92(3): 957–982.

Boggs RVO (1998) The future of civil rights testing: Current trends and new directions. In: Fix M and Turner MA (eds) *A National Report Card on Discrimination in America: The Role of Testing*. Washington, DC: The Urban Institute Press.

Boggs RVO, Sellers JM and Bendick Jr M (1993) Use of testing in civil rights enforcement. In: Fix M and Struyk R (eds) *Clear and Convincing Evidence: Measurement of Discrimination in America*. Washington, DC: The Urban Institute Press.

Bovenkerk F (1992) *Testing Discrimination in Natural Experiments: A Manual for International Comparative Research on Discrimination on the Grounds of 'Race' and Ethnic Origin*. Geneva: International Labour Office.

British Sociological Association (2002) *Statement of Ethical Practice for the British Sociological Association*. Available at: http://www.britsoc.co.uk/media/27107/StatementofEthicalPractice.pdf (accessed 22 September 2015).

Bursell M (2007) What's in a name? A field experiment test for the existence of ethnic discrimination in the hiring process. *SULCIS Working Paper*. Available at: www.temaasyl.se/Documents/Forskning/…/What's%20in%20a%20name.pdf (accessed 10 September 2014).

Calvès G (2007) Au service de la connaissance et du droit: le testing. *Horizons Stratégiques* 3(5): 8–16.

Carlsson M and Rooth D-O (2007) Evidence of ethnic discrimination in the Swedish labour market using experimental data. *Labour Economics* 14(1): 716–729.

Carlsson M and Rooth D-O (2012) The power of media and changes in discriminatory behaviour among employers. *Journal of Media Economics* 25(2): 98–108.

Cherry F and Bendick M (2018) Making it count: Discrimination auditing and the activist scholar tradition. In: *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. New York: Springer, pp.45–62.

Connor JM (2000) Studying racial bias: Too hot to handle? *Journal of College Science Teaching* 30(1): 26–32.

Crabtree C (2018) An introduction to conducting email audit studies. In: *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. New York: Springer, pp.103–117.

Dench S, Iphofen R and Huws U (2004) *An EU Code of Ethics for Socio-Economic Research*. The Institute for Employment Studies. Available at: https://www.employment-studies.co.uk/ (accessed 8 September 2015).

Deutsche Gesellschaft für Soziologie (2017) *Ethik-Kodex der deutschen Gesellschaft für Soziologie (DSG) und des Berufsverbandes deutscher Soziologinnen und Soziologen (BDS)*. Available at: https://www.soziologie.de/de/die-dgs/ethik/ethik-kodex/ (accessed 18 August 2018).

Dingwall R (2012) How did we ever get into this mess? The rise of ethical regulation in the social sciences. In: Love K (ed) *Ethics in Social Research*. Bingley: Emerald Group, pp.3–26.

Edley Jr. C (1993) Implications of empirical studies on race and discrimination. In: Fix M and Struyk R (eds) *Clear and Convincing Evidence: Measurement of Discrimination in America*. Washington, DC: The Urban Institute, pp.377–391.

Eid P (2012) *Mesurer la Discrimination à l'Embauche Subie par les Minorités Racisées: Résultats d'un 'Testing' Mené dans le Grand Montréal*. Montréal: Commission des Droits de la Personne et des Droits de la Jeunesse.

Fix M, Galster G and Struyk R (1993) An overview of auditing for discrimination. In: Fix M and Struyk R (eds) *Clear and Convincing Evidence: Measurement of Discrimination in America*. Washington DC: The Urban Institute, pp.1–49.

Gaddis SM (2018a) *Audit Studies: Behind the Scenes with Theory, Method, and Nuance. Methods Series*. Cham: Springer.

Gaddis SM (2018b) An introduction to audit studies in the social sciences. In: *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Cham: Springer, pp.3–44.

Gingrich N (1998) Testimony of House Speaker Newt Gingrich before the House Subcommittee on Employer-Employee Relations on 'The Future Direction of the Equal Employment Opportunity Commission'. Available at: archives-republicans-edlabor.house.gov/archive/hearings/105th/eer/…/gingrich.htm (accessed 17 October 2014).

Goldberg A, Mourinho D and Kulke U (1995) *Labour Market Discrimination against Foreign Workers in Germany*. International Labour Office, Employment Department. Available at: https://www.ilo.org/wcmsp5/groups/public/—…/—migrant/…/wcms_201039.pdf (accessed 1 October 2014).

Google (2014) *Google Terms of Service (Last updated April 14, 2014)*. Available at: http://www.google.com/intl/en/policies/terms/.

Hunter D (2010) Locating ethics in research. In: Hughes J (ed) *European Textbook on Ethics in Research*. Luxembourg: Publication Office of the European Union, pp.11–32.

Israel M (2015) *Research Ethics for Social Scientists*. London: Sage.

Jowell R and Prescott-Clarke P (1970) Racial discrimination and white-collar workers in Britain. *Race & Class* 11(4): 397–417.

Kaas L and Manger C (2012) Ethnic discrimination in Germany's labour market: A field experiment. *German Economic Review* 13(1): 1–20.

Klose A and Kühn K (2009) *Gewerberecht, Rassistische/Ethnische Diskriminierung und Testingverfahren*. Berlin: Antidiskriminierungsnetzwerks Berlin des Türkischen Bundes in Berlin-Brandenburg.

Klose A and Kühn K (2010) Die Anwendbarkeit von Testing-Verfahren im Rahmen der Beweislast, § 22 Allgemeines Gleichbehandlungsgesetz. *Expertise im Auftrag der Antidiskriminierungsstelle des Bundes, Berlin*. Available at: cet.lu/wp-content/uploads/2012/10/KLOSE.ppt (accessed 5 May 2015).

Koopmans R, Veit S and Yemane R (2018) Ethnische Hierarchien in der Bewerberauswahl: Ein Feldexperiment zu den Ursachen von Arbeitsmarktdiskriminierung. *WZB Discussion Paper* SP VI 2018-104. Available at: https://bibliothek.wzb.eu/pdf/2018/vi18-104.pdf, (accessed 18 August 2018).

Kühn K, Liebscher D and Klose A (2013) Expertise zu den rechtlichen Aspekten sozialwissenschaftlicher Testing-Verfahren im Rahmen der vom SVR geplanten Studie 'Diskriminierung am Ausbildungsmarkt?' Available at: http://svr-migration.dymagics.de/wpcontent/uploads/2014/11/Expertise_fuer_Studie_Diskriminierung_am_Ausbildungsmarkt.pdf (accessed 4 September 2015).

Lahey J and Beasley R (2018) Technical aspects of correspondence studies. In: Gaddis M (ed) *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Cham: Springer, pp.81–101.

LaPiere RT (1934) Attitudes vs. actions. *Social Forces* 13(2): 230–237.

Larja L, Warius J and Sundbäck L et al. (2012) *Discrimination in the Finnish Labour Market: An Overview and a Field Experiment on Recruitment*. Helsinki: Ministry of Employment and the Economy – Employment and Entrepreneurship, European Commission.

Liebkind K, Larja L and Brylka A (2016) Ethnic and gender discrimination in recruitment: Experimental evidence from Finland. *Journal of Social and Political Psychology* 4(1): 403–426.

Microsoft (2015) *Microsoft Services Agreement (Published June 4, 2015)*. Available at: https://www.microsoft.com/en-us/servicesagreement/.

Midtbøen AH (2013) *Determining Discrimination*. University of Oslo.

Midtbøen AH (2014) The invisible second generation? Statistical discrimination and immigrant stereotypes in employment processes in Norway. *Journal of Ethnic and Migration Studies* 40(10): 1657–1675.

Nakray K (2016) Social science research ethics for a globalising world – A critical overview of interdisciplinary and cross-cultural perspectives. In: Nakray K, Alston M and Whittenbury K (eds) *Social Science Research Ethics for a Globalizing World – Interdisciplinary and Cross-Cultural Perspectives*. New York: Routledge.

National Committees for Research Ethics in Norway (NESH) (2006) *Guidelines for Research Ethics in the Social Sciences, Law and the Humanities*. Available at: https://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/guidelines-for-research-ethics-in-the-social-sciences-law-and-the-humanities-2006.pdf (accessed 15 January 2015).

Neumark D (2016) Experimental Research on Labour Market Discrimination. *NBER Working Paper No. 22022*. Available at: https://ideas.repec.org/p/nbr/nberwo/22022.html (accessed 7 March 2016).

Neumark D, Burn I and Button P (2017) Age discrimination and hiring of older workers. In: *FRBSF Economic Letter*. Federal Reserve Bank of San Francisco.

Pager D (2007) The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *Annals of the American Academy of Political and Social Science* 609(1): 104–133.

Pager D and Western B (2012) Identifying discrimination at work: The use of field experiments. *Journal of Social Issues* 68(2): 221–237.

Quillian L, Pager D and Hexel O et al. (2017) Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences* 114(41): 10870–10875.

Riach PA and Rich J (2004a) Deceptive field experiments of discrimination: Are they ethical? *Kyklos* 57(3): 457–470.

Riach PA and Rich J (2004b) Fishing for discrimination. *Review of Social Economy* 62(4): 465–486.

Rich J (2014) What do field experiments of discrimination in markets tell us? A meta analysis of studies conducted since 2000. *IZA Discussion Paper*. Available at: ftp.iza.org/dp8584.pdf. (accessed 19 September 2018).

Rorive I (2009) *Proving Discrimination Cases: The Role of Situation Testing*. Bruxelles and Stockholm: Centre for Equal Rights.

Schneider J, Yemane R and Weinmann M (2014) *Diskriminierung am Ausbildungsmarkt: Ausmass, Ursachen und Handlungsperspektiven*. Berlin: Forschungsbereich beim Sachverständigenrat Deutscher Stiftungen für Integration und Migration (SVR).

Simitian J (2010) An act to add Section 528.5 to the Penal Code, relating to impersonation. In: California So (ed) *SB 1411*. Available at: http://leginfo.ca.gov/pub/09-10/bill/sen/sb_1401-1450/sb_1411_bill_20100927_chaptered.html (accessed 18 February 2016).

van der Plancke V (2007) Les tribulations du testing en Belgique: Quels enseignements? *Horizons Stratégiques* 3(5): 40–59.

Vuolo M, Uggen C and Lageson S (2016) Statistical power in experimental audit studies: Cautions and calculations for matched tests with nominal outcomes. *Sociological Methods & Research* 45(2): 260–303.

Vuolo M, Uggen C and Lageson S (2018) To match or not to match? Statistical and substantive considerations in audit design and analysis. *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Cham: Springer, pp.119–140.

Wassenaar D (2006) Ethical issues in social science research. In: Terre Blanche M, Durrheim K and Painter D (eds) *Research in Practice: Applied methods for the Social Sciences*. 2nd ed. Capetown: University of Capetown Press.

Weichselbaumer D (2015) Testing for discrimination against lesbians of different marital status: A field experiment. *Industrial Relations: A Journal of Economy and Society* 54(1): 131–161.

Weichselbaumer D (2016) Discrimination against female migrants wearing headscarves. *IZA Discussion Paper Series*. Available at: http://ftp.iza.org/dp10217.pdf (accessed 21 June 2018).

Wood M, Hales J and Purdon S et al. (2009) *A Test for Racial Discrimination in Recruitment Practice in British Cities*. Norwich: Department for Work and Pensions.

Yahoo (2012) *Yahoo Terms of Service (Last updated March 16, 2012)*. Available at: https://policies.yahoo.com/us/en/yahoo/terms/utos/.

Zschirnt E (2016) Measuring hiring discrimination – A history of field experiments in discrimination research. *NCCR Working Papers* 1–32. Available at: https://nccr-onthemove.ch/publications/measuring-hiring-discrimination-a-history-of-field-experiments-in-discrimination-research/ (accessed 15 June 2016).

Zschirnt E and Ruedin D (2016) Ethnic discrimination in hiring decisions: A meta-analysis of correspondence tests 1990–2015. *Journal of Ethnic and Migration Studies* 42(7): 1115–1134.